

Advanced Programming Guide with C++ XML

C++XML

高级编程指南

[美] Fabio Arciniegas 著
武磊 刘拴强 译



New
Riders

中国科学院出版集团
北京希望电子出版社

Advanced Programming Guide with C++ XML

C++XML

高级编程指南

[美] Fabio Arciniegas 著
武磊 刘拴强 译



New
Riders

中国科学出版集团
北京希望电子出版社

内 容 简 介

本书写给那些专门应用 XML 和 C++ 开发高性能、高质量应用程序的程序员。在被初级入门信息充塞的市场中, 这本书为成功建构 C++ XML 程序提供了有效、高水平 and 正中要点的技巧和工具, 这些程序的范围涵盖了从 XML 编辑器和数据库到 3D 应用程序和分布式系统的所有内容。

全书由 5 部分构成。内容分别为: XML 结构和概念, XML/C++ 概述, 事件驱动过程, SAX C++, SAX C++ 2.0 及高级技术, DOM Level 2.0, C++ DOM 高级操作, SAX 和 DOM 的高级 C++ 特征, XML 模式、TREX 和其他建模语言, C++ 中的 Xpath 和 XPointer, XSLT 转换, 工具包特性和框架实现, 为 C++ 程序创建基于 XML 的扩展语言, 已发行 C++ XML 应用程序: SOAP 和 XML-RPC, 使用 C++ 开发 XML 和数据库 (ODBC 和 DAO), UML, STL, CD 内容和说明。

本教程详尽完整, 专业性强, 是那些应用 XML 和 C++ 构建应用程序的程序员首选参考书, 同时本书也可作为高等院校相关专业师生教学、自学参考书。

本版 CD 内容为书中的实例源代码及可执行文件。

版 权 声 明

本书英文名为 “C++ XML”, 由 New Riders Publishing 出版, 版权归 New Riders Publishing 所有。本书中文版由 New Riders Publishing 授权出版。未经出版者书面许可, 本书的任何部分不得以任何手段复制和传播。

盘 书 系 列 名 : 21 世纪新工具软件开发指南丛书 (6)

盘 书 名 : Advanced Programming Guide with C++ XML C++XML 高级编程指南

文 本 著 作 者 : (美) Fabio Arciniegas 著 武磊 刘拴强 译

责 任 编 辑 : 王玉玲

CD 制 作 者 : 希望多媒体开发中心

CD 测 试 者 : 希望多媒体测试部

出版、发 行 者 : 北京希望电子出版社

地 址 : 北京中关村大街 26 号, 100080

网址: www.bhp.com.cn E-mail: lwm@hope.com.cn

电话: 010-62562329, 62541992, 62637101, 62637102, 62633308, 62633309 (发行部)

010-62613322-215 (门市) 010-62547735 (编辑部)

经 销 : 各地新华书店、软件连锁店

排 版 : 希望图书输出中心 董淑红

CD 生 产 者 : 北京中新联光盘有限责任公司

文 本 印 刷 者 : 北京媛明印刷厂

开 本 / 规 格 : 787 毫米×1092 毫米 16 开本 19.75 印张 347 千字

版 次 / 印 次 : 2002 年 3 月第 1 版 2002 年 3 月第 1 次印刷

印 数 : 1-4000 册

本 版 号 : ISBN 7-900088-52-0

定 价 : 35.00 元 (本版 CD)

说明: 凡我社产品如有残缺, 可执相关凭证与本社调换。

关于作者

Fabio Arciniegas 是 postgraphy (<http://www.postgraphy.com>) 的首席技术专家, 他将 C++ 和 XML 技术应用到了开发可视和图像处理应用中。他是 XML 协会中的活跃的成员之一, 不仅是开发者, 也是作者。他的贡献包括了 C++ 实现 TREX 语言的开发, Xlink 软件, 大量的文章(尤其是在 www.xml.com 上的)以及处理 XML 技术的指导程序。他经常作为 XML 相关会议的发言人和指导者参与活动, 包括了 XML Developers Conference, O'Reilly Open Source Conference, 以及 XML and Distributed Applications (他现在主席) 等。

关于技术编辑

在本书写作的过程中, 技术编辑审阅了所有的内容。

Beth Breidenbach 是 Getronics 的产品设计师, 提供软件和基础结构的解决方案。自称为 “data geek” 的她, 有着对所有数据设计、存储、转换和翻译等问题的持久兴趣。她现在正在研究自然语言界面的数据设计问题。

Jason A. Buss 很早就开始从事计算机工作。在做了 CNC 的程序员之后, 他在 1996 年进入了技术出版领域。现在他正在为一家大型飞机制造商开发和改进 SGML/XML 出版系统。他的兴趣包括组装电脑, 开发开放源代码的 XML 软件, 阅读和弹吉他。

译者的话

在 Internet 风靡全球、Web 技术日趋成熟的今天，“数字地球”、“数字家园”、“电子商务”等网络相关技术已经走入我们的生活。此时，传统的 HTML、DHTML 等语言标准由于其整合数据的能力太差而限制了网络技术的深入发展。1998 年 2 月 10 日，W3C（万维网协会）成员鉴定和批准了 XML 作为新的网络语言推荐标准，对于广大网络程序员来说，这无疑是给他们注入了一支兴奋剂。XML（eXtensible Markup Language，可扩展标记语言）是 SGML 语言（标准通用标记语言）的一个子集，由于它有着较大的灵活性、更强的整合数据能力，很快便为广大程序员所喜爱。有人甚至认为它给目前的 Web 技术带来一次革命，是“第二代 Web”。因此，能将这本 XML 的专业书籍奉献给大家，我们感到由衷的高兴。

本书是学习 XML 较为专业的书籍。不同于一般 XML 书籍的是，本书主要讨论如何用 C++ 语言实现 XML 的高级技术，而且它的内容起点比较高，适合于对 XML 技术有一定认识并有一定 C++ 编程经验的网络编程人员。如果读者以前从没有接触过 XML 但现在对它感兴趣的话，我们建议您先阅读相关的基础书籍。本书所附的 CD 光盘中包含书中所有程序的代码，其中大多数程序都是在 Windows Visual Studio 6.0 环境下开发的，但也有部分程序是基于 Linux 平台的。所以除了 Windows 外，读者还应该有一个运行 Linux 程序的环境。

本书第一章介绍了有关 XML 的一些基本结构和概念，如果读者对 XML 已经比较熟悉的话可以跳过此章。相反，如果是刚刚接触 XML，可能需要在这一章多花些功夫。第二章和第三章主要讲述如何用 C++ 处理 XML 以及二者之间的关系，这些是读者深入学习用 C++ 处理 XML 应用的基础。第四章和第五章分别是对 SAX（XML 简单 API）的介绍和 SAX2.0 的接口特性的说明。第六章、第七章讲述有关文档对象模型（DOM）的知识，第八章则比较了 C++ DOM 和 C++ SAX，并介绍了工具集 MSXML 和 Xerces。从第九章到第十一章，介绍 XML 家族的一些主要部件，如 XML 模式、XPath、XPointer 和 XSLT 等，第十二章则讲了如何运用这些工具实现常见的任务。第十三章是关于如何为 C++ 程序开发基于 XML 的扩展语言，第十四章介绍已发行的 C++ XML 应用程序：SOAP 和 XML-RPC，展示了用 C++ 实现应用程序中高性能的组件的机制。本书的最后一章——第十五章，作者重点介绍了使用 C++ 开发 XML 及数据库（ODBC 和 ADO）的相关技术，这也是 C++ XML 编程最激动人心的地方。附录提供了有关统一建模语言（UML）和标准模板库（STL）的知识，并介绍了本书所附 CD 的内容。

本书由刘拴强（翻译第一、三、五、七、九、十一、十三、十五章）和武磊（翻译第二、四、六、八、十、十二、十四章及其它部分）共同完成翻译，并由杨瑞完成校对。本书的翻译得到了希望出版社刘晓融女士的热心帮助和指导，我们在此谨表衷心感谢。

限于水平，本书的翻译缺漏和错误之处在所难免，恳请广大读者不吝指正。

译者

2005/29/07

引 言

本书写给那些专门应用 XML 和 C++ 开发高性能、高质量应用程序的程序员。在被初级入门信息充塞的市场中，这本书为成功建构 C++ XML 程序提供了有效、高水平 and 正中要点的技巧和工具，这些程序的范围涵盖了从 XML 编辑器和数据库到 3D 应用程序和分布式系统的所有内容。

本书涵盖内容

从面向技能的观点来看，本书涵盖了以下内容：

- 关于有效的服务器端、单机和客户端为 XML 使用 C++ 开发的所有方面。
- 用于 Windows 和 Linux XML 应用程序开发的高级技术。
- 所有应用主流 XML C++ 工具和框架的重要细节。
- 跨平台的集成和扩展技术。
- 应用 XML 和 C++ 开发分布式应用程序的最新信息。

从技术角度看，本书涵盖了以下内容：

- XML 1.0+ 命名空间
- SAX 2.0
- DOM 2.0
- Non-SAX/DOM 解析器
- XPath
- XPointer
- XML Schema 和其他建模语言
- C++ 的编码
- XML 和数据库
- XSLT
- 使用 XML 作为 C++ 应用程序的扩展语言
- 使用 C++ 独有特性（例如模板）作为 XML 处理过程扩展模型
- 服务器方面的话题，例如使用 C++ 集成 XML+XSLT+Apache。

本书使用对象

本书的对象是那些应用 XML 和 C++ 构建应用程序的程序员。

如果读者需要一本使用 C++ 阐明操作所有主要的 XML 技术的书，那么本书是您明智的选择。

无论读者是想开发诸如 OpenGL SAX 之类将 XML 作为其 3D 文件格式的应用程序的软件、一个应用 SOAP 或 XML-RPC 消息的 Windows MFC 应用程序、一个应用 DOM 的 Linux GTK 应用程序、一个使用 ADO 和 ODBC 数据库的 XML 视图的 Windows 应用程序、使用 C++ XSLT 的扩展语言解释程序，还是其他任何 C++ XML 应用，本书都会提供相关理论、代码和生成它所需的内部信息。

这是一本有关 C++ 的书。如果需要本书所讨论的 XML 主题及其他方式的说明，读者可参阅一些通用的书，例如“Inside XML”(New Riders, 2001)或者“XML Developer's Guide”(McGraw-Hill, 2000)。

本书定位和设计原则

本书定位在中高级之间，遵循如下著写原则：

- 详尽。本书涉及了创建非凡的 XML 应用程序的所有相关技术、方法和必须的技巧。（请参阅第二章以得到完整的组织结构）。
- 支持理论的绝对代码。那些代码支持每个 API、结构和方法。本书所示的每一行代码都是一个完全可用的 C++ 程序的一部分，程序和完整的源代码包含在 CD 中。
- 可重用性。本书包括的程序、模块和模式已经写成尽可能在读者自己的工程中得以重用的形式。

使用本书先决条件

从读者的观点来看，惟一需要的就是要熟悉 C++。稍微了解一些关于 XML 格式也是很有帮助的（尽管第一章对 XML 1.0+ 命名空间的所有概念做了详尽的解释）。

所需的硬件和软件是一台个人电脑和一个 C/C++ 编译器。书中的实例主要是在 Windows Visual Studio 6.0 环境下开发的，大多数可执行程序都是 Win32 应用程序。我们也给出了 ELF 和本书包括的 Linux 应用程序。还包含了使用 gcc 2.91 或更高版本的 Linux 文件和 Windows 的 Cygnus gcc 2.95 或更高版本。

致谢

对 Stephanie、Laura Loveall、Allison Johnson、Bech Breidenbach、Jason Buss 和 New Riders 等所有 New Riders 工作组致以特别的感谢。最后还要感谢我的家人和那些对我的困惑给予指点的朋友，他们使得本书的内容真臻于完善。

把您的想法告诉我们

作为本书的读者，您是最重要的评论者和批评者。我们非常重视您的观点并想知道我

们做的哪些是正确的，怎样才能做得更好，您还希望我们介绍哪些方面的内容，以及其他您想要告诉我们的精辟见解。

作为 New Riders Publishing Web 开发组的执行主编，我欢迎您提出意见。您可以通过传真、电子邮件或者直接给我写信来让我知道对本书您喜欢什么、不喜欢什么——以及我们怎样做才能使得本书更加完美。

请注意我不能帮助您解决与本书主题相关的技术问题，这取决于我所收到信件的数量，我可能无法回复每条消息。

您在写信时请务必包含本书的题目和作者，以及您的名字和电话号码或者传真号码。我会仔细考虑您的意见，并与本书的其他作者和编辑共享它们。

Fax: 317-581-4663

Email: Stephanie.wall@newriders.com

Mail: Stephanie Wall

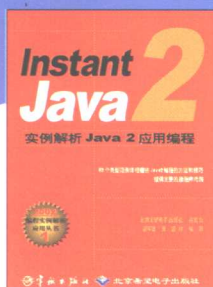
Executive Editor

New Riders Publishing

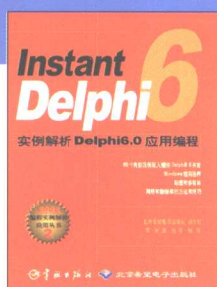
201 West 103rd Street

Indianapolis, IN 46290 USA

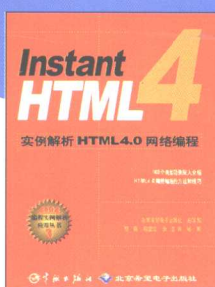
专业程序员成长之路



CX-3567
定价: 40.00 元



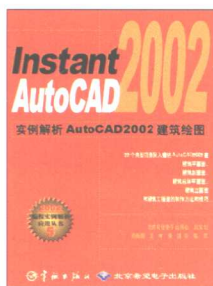
CX-3568
定价: 40.00 元



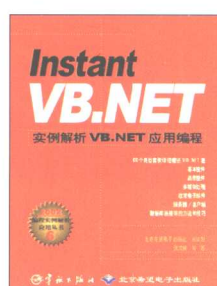
CX-3571
定价: 40.00 元



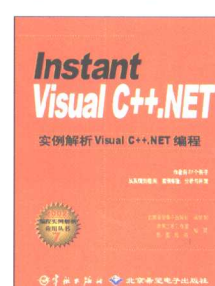
CX-3572
定价: 40.00 元



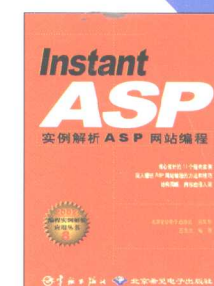
CX-3569
定价: 40.00 元



CX-3581
定价: 40.00 元



CX-3583
定价: 40.00 元



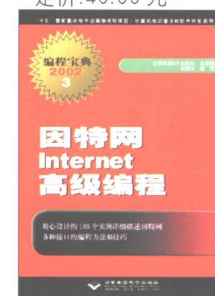
CX-3591
定价: 40.00 元



CX-83576
定价: 58.00 元



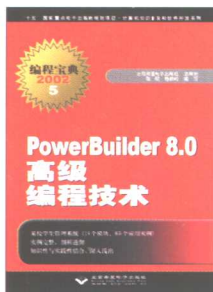
CX-83598
定价: 30.00 元



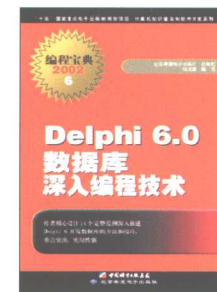
CX-83597
定价: 55.00 元



CX-83599
定价: 35.00 元



CX-83620
定价: 43.00 元



CX-83611
定价: 35.00 元



CX-83613
定价: 35.00 元



CX-83618
定价: 33.00 元



北京希望电子出版社
Beijing Hope Electronic Press
www.bhp.com.cn

地址: 北京中关村 083 信箱北京希望电子出版社 (邮编 100080)
电话: 010-62562329、010-62633308 传真: 010-62579874
E-mail: qrh@hope.com.cn

专业程序员成长的摇篮



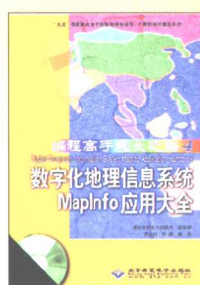
CX-83533
定价: 46.00 元



CX-83512
定价: 42.00 元



CX-83513
定价: 55.00 元



CX-83480
定价: 66.00 元



CX-83442
定价: 50.00 元



CX-83407
定价: 46.00 元



CX-83391
定价: 55.00 元



CX-83558
定价: 58 元



CX-83549
定价: 35 元



CX-82893
定价: 58 元



CX-82892
定价: 46 元



CX-83203
定价: 62.00 元



CX-3317
定价: 58.00 元



CX-3318
定价: 62.00 元



CX-3316
定价: 38.00 元



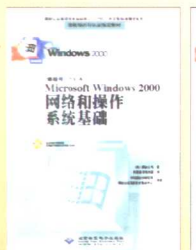
CX-3314
定价: 58.00 元



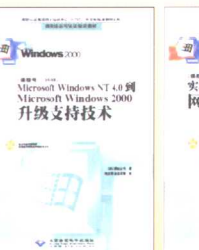
CX-3315
定价: 38.00 元



CX-83279
定价: 55.00 元



CX-83205
定价: 42 元



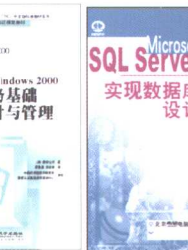
CX-83213
定价: 55 元



CX-83215
定价: 66 元



CX-83181
定价: 30 元



CX-83219
定价: 58.00 元



CX-2660
定价: 42 元



CX-2663
定价: 44 元



北京希望电子出版社
Beijing Hope Electronic Press
www.bhp.cn

社址: 北京中关村大街26号(黄庄路口东) 通讯: 北京中关村083信箱(100080)
电话: (010) 62562329 62541992 传真: (010) 62579874 62633308

目 录

I XML 基础

第一章 XML: 结构和概念	3	1.3.2 声明	18
1.1 基本结构	3	1.3.3 通用名称和受限名称	19
1.1.1 元素、属性、和嵌套	3	1.3.4 在 DTD 中声明命名空间	19
1.1.2 XML 声明	4	1.4 总结	21
1.1.3 CDATA 部分	5	第二章 XML/C++概览	22
1.1.4 处理指令	5	2.1 XML 技术家族	22
1.1.5 自然语言	5	2.1.1 核心 XML	22
1.1.6 格式正规	6	2.1.2 高级数据建模和验证	24
1.2 DTDs	6	2.1.3 定位和链接技术	27
1.2.1 元素类型声明	6	2.1.4 XML 文档转换	29
1.2.2 对 DTD 绑定文档	8	2.2 APIs	31
1.2.3 属性列表声明	8	2.2.1 通用基础 API	31
1.2.4 表示法	12	2.2.2 特定词汇表 API	33
1.2.5 稍作休息: 关于 “show DTD”	12	2.3 应用程序	33
1.2.6 条件部分	15	2.3.1 纯 XML 操作	34
1.2.7 实体	15	2.3.2 相关的补充	34
1.2.8 字符引用	17	2.3.3 分布式 XML 应用程序	34
1.2.9 内部和外部 DTD 子集	17	2.3.4 案例学习	35
1.3 命名空间	18	2.4 总结	35
1.3.1 基本原理	18		

II 用 C++处理 XML

第三章 事件驱动过程	39	3.2.7 expat 关于面向对象的封装	49
3.1 事件驱动模式	39	3.2.8 在 Windows 工程中集成 expat	
3.2 Expat	40	(可选的)	53
3.2.1 实地介绍	40	3.3 其它 C/C++工具箱	54
3.2.2 语法树类	41	3.4 总结	57
3.2.3 处理器	44	第四章 SAX C++	58
3.2.4 结果	47	4.1 实地介绍	59
3.2.5 expat 的结构	47	4.1.1 文档类型	59
3.2.6 基本函数	47	4.1.2 策略	60

4.2 SAX C++的结构	66
4.3 SAX 1.0	66
4.3.1 处理器	67
4.4 总结	81
第五章 SAX C++2.0 及高级技术	82
5.1 通用接口	82
5.1.1 DTDHandler	82
5.1.2 ErrorHandler	83
5.1.3 使用 ErrorHandler	84
5.1.4 基础处理器	85
5.2 SAX 2.0 的专有接口和类	85
5.2.1 SAX 2.0 和 SAX 1.0 的主要不同	85
5.2.2 有关 SAX2 的一张图	86
5.2.3 源头: SAX2XMLReader	90
5.2.4 SAX2XMLReader 方法	91
5.2.5 Reader 工厂	97
5.2.6 助手和异常	97
5.3 总结	98
第六章 DOM Level 2.0	99
6.1 DOM 基础	99
6.2 DOM 发展历史	100
6.3 DOM 性质	100
6.3.1 结构	101
6.4 一个实例: XML 编辑器	104
6.5 核心接口索引	108

6.6 一个完整的例子: XML 编辑器	108
6.7 总结	114
第七章 C++ DOM 高级操作	115
7.1 将文档模式绑定到本地模式	115
7.1.1 对问题的定义	115
7.1.2 问题	116
7.1.3 模型和视图与模型和结构	119
7.1.4 归纳方法	120
7.2 用检查者扩展 DOM	120
7.2.1 统计检查者	121
7.2.2 命名空间检查者	124
7.3 总结和下一步的任务	125
第八章 SAX 和 DOM 的高级 C++特征 ...	126
8.1 C++ SAX 和 C++ DOM	126
8.1.1 性能	126
8.1.2 关于设计	128
8.1.3 关于模式	130
8.2 C++ SAX 加 C++ DOM	130
8.2.1 DOMAsSAX 适配器	130
8.3 C++ SAX 中的内存管理	135
8.3.1 字符串的分配和释放	135
8.4 C++ DOM 中的内存管理	136
8.4.1 C++中的字符编码	137
8.5 总结	143

III 使用与 C++相关的 XML 标准

第九章 XML 模式、TREG 和其它

建模语言	147
9.1 XML 模式: 语言概述	147
9.1.1 XML 模式的基本原理	147
9.1.2 类型	147
9.1.3 复杂类型与简单类型	149
9.1.4 匿名类型	151
9.1.5 目标命名空间	151

9.1.6 类型约束: 各方面	152
9.1.7 关于 XML 模式语言的更多内容	152
9.2 XML 模式的 C++工具	153
9.2.1 在 C++工程中嵌入模式验证器	153
9.2.2 存取 XML 模式信息	155
9.3 TREG: 语言概述	155
9.3.1 模式	156
9.3.2 元素和属性	157

9.3.3 选择与循环.....	158
9.3.4 交叉.....	160
9.4 用于 TREX 的 C++ 工具: Lorenza	161
9.4.1 作为应用程序运行 Lorenza	162
9.4.2 嵌入 Lorenza	162
9.4.3 有计划地得到模式结构信息	164
9.4.4 RELAX 和 Schematron.....	164
9.5 总结	166
第十章 C++ 中的 XPath 和 XPointer ..	167
10.1 举例讲解 XPath.....	167
10.1.1 文档.....	167
10.1.2 表达式及其值	171
10.2 正式开始 XPath.....	173
10.2.1 模型: 轴和节点	173
10.2.2 谓词.....	175
10.2.3 函数.....	175
10.3 Xpath 的 C++ 操作.....	177
10.3.1 Xalan 中的 XPath 类	177
10.3.2 XPath 类: 角色和交互	178
10.3.3 实现 XPath 交互	178
10.4 XPointer	184
10.4.1 点.....	184
10.4.2 范围.....	184
10.4.3 定位集.....	184
10.4.4 附加函数.....	184
10.4.5 语法修饰.....	185
10.5 XPointer 的 C++ 操作	186
10.6 XPointer 类.....	186
10.6.1 图形小说浏览器	187
10.7 总结	191
第十一章 XSLT 转换	192

11.1 XSLT	192
11.1.1 转换过程.....	192
11.1.2 源文档.....	194
11.1.3 预期结果.....	196
11.1.4 在结果中插入数据	197
11.1.5 控制流.....	198
11.1.6 选择.....	200
11.1.7 循环.....	201
11.2 C++ 中的 XSLT	202
11.2.1 “低级” API	203
11.2.2 “高级” API	203
11.2.3 示例应用程序: vxslt	204
11.2.4 更多的例子.....	207
11.3 总结	208
第十二章 工具包特性和框架实现 ...	209
12.1 Xerces/Xalan	209
12.1.1 平台和许可证.....	209
12.1.2 支持技术	210
12.1.3 哪里可以得到它.....	210
12.1.4 什么是框架.....	210
12.1.5 XMLableFR (XMLable 框架)	211
12.1.6 其他可高度重用的 Xerces/Xalan 软件	224
12.2 MSXML	224
12.2.1 平台和许可证.....	224
12.2.2 支持技术	225
12.2.3 哪里可以得到它.....	225
12.2.4 可重用的设计和代码: DOM Walkers	225
12.2.5 实现 DOMWalkerWizard	233
12.3 总结	239

IV 集成著名的 C++ 模块

第十三章 为 C++ 程序创建基于 XML 的扩展语言	243
--	------------

13.1 从 XML 创建程序树对象	243
13.1.1 机理总览	244

13.1.2	语言的设计	244
13.1.3	基本原理	245
13.1.4	基本要素	245
13.1.5	变量和操作符	247
13.1.6	控制结构	248
13.1.7	创建对象结构	251
13.1.8	基本要素建模	251
13.1.9	控制结构建模	252
13.1.10	构造 Term 树	253
13.2	总结	253
第十四章 已发行 C++ XML 应用程序:		
	SOAP 和 XML-RPC	254
14.1	对问题的介绍	254
14.2	SOAP 1.1	254
14.2.1	封套	254
14.2.2	SOAP 编码	256
14.2.3	RPC 约定	257
14.3	SOAP HTTP	259

14.3.1	MS C++ SOAP 实现	259
14.4	XML-RPC	262
14.4.1	请求和反馈	262
14.4.2	错误	264
14.5	独立 C++ 计算器服务	265
14.6	独立的 C++ 客户端	266
14.7	总结	267

第十五章 使用 C++ 开发 XML 和数据库

	(ODBC 和 DAO)	268
15.1	XML 在数据库描述中的角色	268
15.2	C++ 工具选项	268
15.2.1	中间件	269
15.3	适配器 APIs	270
15.3.1	类特定伙伴	271
15.3.2	技术特定 XML 支持	276
15.3.3	外部 C++ 数据库资源	279
15.4	总结	279

V 附录

附录 A UML		283
A.1	类图	283
A.1.1	类	283
A.1.2	接口	286
A.1.3	关系	286
A.2	接口图	287
A.2.1	顺序图	288
A.2.2	协作图	289
附录 B STL		290
B.1	STL 性质和组件	290
B.1.1	容器	291
B.1.2	迭代器	293

B.1.3	通用算法	294
B.1.4	函数对象	296
B.1.5	适配器和分配器	298
B.1.6	STL 头	299
附录 C CD 内容和说明		300
C.1	所需条件	300
C.2	平台	300
C.3	组织结构	300
C.4	安装说明	300
C.5	程序更新	301
C.6	有关 Extra-Goodies 文件夹	301
C.7	Splash Screens 和 CSS 安装	301

I

XML 基础

第一章 XML：结构和概念

第二章 XML/C++概览

第一章 XML：结构和概念

本书不是面向所有读者的 XML 入门书（关于这方面的介绍，已经有大量的书籍和网上的免费资源了），相反地，本书浓缩了利用 C++ 语言处理 XML 的高级技术。所以我们假定读者对 C++ 和 XML 已经有了基本的了解。

然而出于完整性的考虑，本章仍然概述了 XML 的结构和概念，范围覆盖了从格式正规到命名空间。如果读者已经对 XML 很熟悉了，那么可以跳过本章，仅把本章作为参考。

本章分为三个部分。第一部分重温了 XML 的基本结构和概念，第二部分是对文档类型定义的回顾，第三部分解释了命名空间。

最后必须要提醒读者注意，本章以 XML 1.0 和命名空间为语言进行讲解，而不涉及 XML 技术家族（以及它们和 C++ 的关系）。那些将在下一章讲解。

1.1 基本结构

我们的讨论从 XML 文档的基本组成部分——元素和属性开始。下一节将讲解其结构和用法。

1.1.1 元素、属性、和嵌套

所有的 XML 文档都是由元素组成的。元素由起始和结束标记界定，可以包含文本或者其他元素。起始标记由标记名和一组属性组成。结束标记仅仅由引导的元素名构成。示例见清单 1.1。

清单 1.1 元素

```
<show>
  <name><simpleName>Tymon Phonty</simpleName></name>
  <episode>
    <description>A Man speaks in anagrams.
    <note> they are really only spoonerisms </note>
    </description>
  </episode>
</show>
```

方便起见，如果一个标记没有内容，则可以简写。例如：

```
<russian> </russian>
```

可简写成：

```
<russian/>
```